

บทที่ 1

บทนำ

1.2 ความสำคัญและที่มา

ในปัจจุบันหลักสูตรระบบสารสนเทศทางธุรกิจคณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เป็นหลักสูตรที่เปิดสอนให้นักศึกษาได้เรียนรู้ความสามารถด้านบริหารธุรกิจ คอมพิวเตอร์และเทคโนโลยีสารสนเทศ มีทักษะในการปฏิบัติงาน ซึ่งโครงสร้างหน่วยกิตของหลักสูตรมีดังนี้

1.1.1 หมวดวิชาศึกษาทั่วไป 30 หน่วยกิต ประกอบด้วยดังนี้

- ภาษาและการสื่อสาร(บังคับ) 12 หน่วยกิต มีจำนวนวิชา 4 วิชา
- สุขภาพ(บังคับ) 3 หน่วยกิต มีจำนวนวิชา 1 วิชา
- บรูณาการ(บังคับ) 9 หน่วยกิต มีจำนวนวิชา 3 วิชา
- สังคมศาสตร์และมนุษยศาสตร์ 3 หน่วยกิต มีจำนวนวิชา 6 วิชา(เลือกเพียงหนึ่งวิชาเท่านั้น)
- วิทยาศาสตร์และคณิตศาสตร์ 3 หน่วยกิต มีจำนวนวิชา 6 วิชา (เลือกเพียงหนึ่งวิชาเท่านั้น)

1.1.2 หมวดวิชาเฉพาะ 103 หน่วยกิต ประกอบด้วยดังนี้

- กลุ่มวิชาพื้นฐานวิชาชีพ 30 หน่วยกิต มีจำนวนวิชา 10 วิชา
- กลุ่มวิชาชีพบังคับ 37 หน่วยกิต มีจำนวนวิชา 13 วิชา
- กลุ่มวิชาชีพเลือก 36 หน่วยกิต มีจำนวน 12 วิชา

1.1.3 หมวดวิชาเลือกเสรี 6 หน่วยกิต มีจำนวน 2 วิชา

รวมหน่วยกิตทั้งสิ้น 139 หน่วยกิต วิชารวมทั้งสิ้น 57 วิชา ทางคณะผู้จัดทำจึงเลือกนำวิชาที่เกี่ยวข้องกับหลักสูตรระบบสารสนเทศมาเป็นข้อมูลในการใช้ ในการวิเคราะห์จากหมวดวิชาศึกษาทั่วไปและหมวดวิชาเฉพาะ มีหน่วยกิตรวมทั้งสิ้น 133 หน่วยกิต มีวิชารวมทั้งสิ้น 47 วิชา

ด้วยการจัดแผนการเรียน โดยไม่มีการวางแผนการเรียนของนักศึกษา หลักสูตรระบบสารสนเทศทางธุรกิจ คณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา โดยไม่มีข้อมูลประกอบการตัดสินใจในการจัดลำดับรายวิชา ทำให้เกิดปัญหาการลงวิชาเรียนในแต่ละปีการศึกษา และไม่ทราบถึงการเรียนการสอนในแต่ละรายวิชาในการศึกษา จึงเกิดปัญหาทำให้

นักศึกษาบางวิชาที่มีความยากและไม่รู้ว่าวิชานั้นๆ มีการเรียนการสอนในรูปแบบใด จึงทำให้มีนักศึกษาติดค้างหรือถอนรายวิชานั้นๆ เป็นจำนวนมาก

จากปัญหาข้างต้นดังนั้นคณะผู้จัดทำจึงนำชุดข้อมูลของ ผลรวมเกรดของแต่ละรายวิชานำมาวิเคราะห์ข้อมูล โดยการใช้ความสามารถในการประมวลผลของโปรแกรม คอมพิวเตอร์เพื่อเพิ่มความน่าเชื่อถือในการวิเคราะห์ และประมวลผลข้อมูลให้เหมาะสมตรงกับความต้องการ โดยการใช้เทคนิคการจัดกลุ่มข้อมูล (Clustering) เป็นกระบวนการสร้างโมเดลจัดการข้อมูล (Data Modeling) ให้อยู่ในกลุ่มที่กำหนดมาให้ โดยการจัดกลุ่มของข้อมูล (Clustering) เมื่อได้ข้อมูลจากการจัดกลุ่มทางผู้จัดทำจะเลือกกลุ่มวิชาที่ยากมาวิเคราะห์พฤติกรรมของนักศึกษาโดยใช้เทคนิคต้นไม้ที่ใช้ในการสนับสนุนการตัดสินใจ (Decision tree) จากนั้นทำการนำเสนอข้อมูล (Data Visualization) ในรูปแบบของภาพด้วยโปรแกรม Power BI และเผยแพร่ข้อมูลสารสนเทศนี้บน เว็บเบราว์เซอร์ (Web Browser) ที่เป็นที่ยอมรับในยุคอินเทอร์เน็ต โดยใช้ภาษา เอกซ์เอ็มแอล (HTML) ซีเอสเอส (CSS) เพื่อเป็นตัวช่วยในการสนับสนุนการตัดสินใจลงทะเลเบียนเรียนและบอกถึงรายละเอียดของแต่ละรายวิชา เพื่อช่วยในการแก้ไขปัญหาจำนวนนักศึกษาที่ติดค้างหรือถอนรายวิชา

1.6 วัตถุประสงค์

เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อการสำเร็จการศึกษาของนักศึกษาโดยใช้วิธีการเหมืองข้อมูล

1.7 ประโยชน์ที่ได้รับ

สามารถทราบปัจจัยที่มีผลต่อการสำเร็จการศึกษาของนักศึกษาโดยใช้วิธีการเหมืองข้อมูล

1.8 ขอบเขต

1.4.1 ขอบเขตผู้วิเคราะห์ข้อมูล

1.4.1.1 Business Understanding ผู้วิเคราะห์ทำความเข้าใจถึงปัญหาและพฤติกรรมของนักศึกษาที่จบแล้ว เพื่อสรุปปัญหาที่ได้และแปลงเป็นแนวคิดให้อยู่ในรูปแบบการวิเคราะห์ดาต้าไมน์นิ่ง

1.4.1.2 Data Understanding วิเคราะห์ข้อมูลเริ่มทำการเก็บรวบรวมข้อมูล และตรวจสอบข้อมูลที่ได้ ดูความถูกต้องและความน่าเชื่อถือของข้อมูล และพิจารณาว่าจะใช้ข้อมูลทั้งหมดทราบปัจจัยที่มีผลต่อการสำเร็จของนักศึกษาโดยใช้วิธีการเหมืองข้อมูล

1.4.1.3 Data Preparation ผู้วิเคราะห์ข้อมูลทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวมมาให้กลายเป็นข้อมูลที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ โดยการทำให้ข้อมูลให้ถูกต้อง

โดยใช้กระบวนการ data cleansing เพื่อให้ได้ข้อมูลอยู่ในรูปแบบและมาตรฐานเดียวกัน และกำจัดข้อมูลเสียออก

1.4.1.4 Modeling คณะผู้จัดวิเคราะห์ข้อมูลด้วยการทำเหมืองข้อมูลแบบ Clustering ที่เอาไว้ใช้การจัดกลุ่มของข้อมูล เพื่อแบ่งกลุ่มวิชาโดยใช้โปรแกรมที่มาใช้ทำเหมืองข้อมูล และนำวิชาที่ยากจากการจัดกลุ่มข้อมูลมาวิเคราะห์พฤติกรรมโดยใช้การทำเหมืองข้อมูลแบบ Decision tree

1.4.1.5 Evaluation ได้ข้อมูลที่ผ่านการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining และการสรุปผลข้อมูลเพื่อให้ตรงกับวัตถุประสงค์ที่ตั้งไว้

1.4.1.6 Deployment นำผลข้อมูลที่ทำการวิเคราะห์แล้ว แสดงผล ข้อมูลบน Web browser โดยใช้ชุดคำสั่ง HTML และ CSS ร่วมกับการนำเสนอข้อมูลแบบ visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพโดยใช้โปรแกรม Power BI ขอบเขตผู้ใช้งานทั่วไปบนเว็บเบราว์เซอร์

1.4.2 ขอบเขตผู้ใช้งาน

1.4.2.1 วิเคราะห์ปัจจัยที่มีผลต่อการเรียนของนักศึกษาหลักสูตรระบบสารสนเทศ จำนวน 47 รายวิชา

1.4.2.2 วิเคราะห์ข้อมูลรายวิชาที่มีผลต่อการสำเร็จการศึกษาด้วยเทคนิคการแบ่งกลุ่มข้อมูล (K-means) และเทคนิคต้นไม้ตัดสินใจ

1.4.2.3 สามารถดูข้อมูลสารสนเทศในลักษณะรูปแบบของแผนภูมิชนิดต่างๆ จากโปรแกรม Power BI

1.4.2.5 สามารถทราบปัจจัยที่ส่งผลต่อการสำเร็จของนักศึกษาผ่าน Web browser

1.9 เครื่องมือที่ใช้ในการพัฒนาโปรแกรม

1.5.1 Hardware

1.5.1.1 คอมพิวเตอร์แล็ปท็อปส่วนตัวเครื่องที่ 1

- CPU – Intel Core i5–8300H
- GPU – NVIDIA GeForce GTX 1050
- RAM – 8GB
- Display – 15.6"
- Storage – HDD 1TB 5400RPM
- OS – Windows 10 Home (64 Bit)

1.5.1.2 คอมพิวเตอร์แล็ปท็อปส่วนตัวเครื่องที่ 2

- CPU – Intel Core i5–8300H
- GPU – NVIDIA GeForce GTX 1050 Ti
- RAM – 8GB
- Display – 16"
- Storage – HDD 1TB .
- OS – Windows 10 Home (64 Bit)

1.5.2 Software

1.5.2.1 โปรแกรม KNIME ใช้ในการทำเทคนิคเหมืองข้อมูล

1.5.2.2 โปรแกรม WEKA ใช้ในการทำเทคนิคเหมืองข้อมูล

1.5.2.3 โปรแกรม Power BI ใช้ในการแสดงผลข้อมูลในรูปแบบของภาพและวิเคราะห์ข้อมูล

1.5.2.4 โปรแกรม MS Excel ใช้จัดรูปแบบของข้อมูล

1.5.2.5 โปรแกรม FileZila โปรแกรมที่ใช้ติดต่อกับ FTP Server เพื่อ อัปโหลด (Upload) ไฟล์ (File)

1.5.2.6 ภาษา HTML ใช้ในการแสดงผลรูปแบบภาพนิทัศน์ของข้อมูล (Data Visualization)

1.5.2.7 ภาษา CSS ใช้ในการตกแต่งเว็บไซต์

1.6 สถานที่ใช้ในการดำเนินการศึกษาและรวบรวมข้อมูล

หลักสูตรระบบสารสนเทศทางธุรกิจคณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ 128 ถนนห้วยแก้ว ตำบลช้างเผือก อำเภอเมือง จังหวัดเชียงใหม่ 50300

1.7 ระยะเวลาในการดำเนินงาน

ตารางที่ 1.1 ระยะเวลาในการดำเนินงาน

แผนการดำเนินการ	2565					
	มิ.ย.	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.
1. ศึกษาและกำหนดความต้องการ	↔					
2. วิเคราะห์ออกแบบระบบและสร้างฐานข้อมูล		↔	↔			
3. เขียนและทดสอบระบบ			↔	↔		
4. วิเคราะห์ข้อมูล (Data Analysis)			↔	↔	↔	↔
4.1 Business Understanding			↔	↔		
4.2 Data Understanding			↔	↔		
4.3 Data Preparation				↔	↔	
4.4 Modeling					↔	↔
4.5 Evaluation					↔	↔
4.6 Deployment						↔
5. แสดงผลการวิเคราะห์ข้อมูลในรูปแบบต่าง ๆ (Data Visualization)					↔	↔
6. นำข้อมูลมาจัดรูปแบบการประชาสัมพันธ์ที่เหมาะสม					↔	↔
7. จัดทำคู่มือการใช้งาน					↔	↔
8. จัดทำเอกสารประกอบโครงการ					↔	↔

1.8 นิยามศัพท์เฉพาะ

1.8.1 การจัดกลุ่มข้อมูล (Clustering) เป็นเทคนิคที่ใช้จำแนกหรือแบ่ง Case หมายถึง คน สัตว์ สิ่งของ หรือ องค์กร ฯลฯ หรือแบ่งตัวแปรออกเป็นกลุ่มย่อย ๆ ตั้งแต่ 2 กลุ่มขึ้นไป Case ที่อยู่ในกลุ่มเดียวกันจะมีลักษณะที่เหมือนกันหรือคล้ายกัน ส่วน Case ที่อยู่ต่างกลุ่มกัน จะมีลักษณะที่แตกต่างกัน ดังนั้น การพิจารณาเลือกลักษณะหรือตัวแปรที่จะนำมาใช้ในการแบ่งกลุ่ม Case จึงมีความสำคัญ

1.8.2 ต้นไม้ตัดสินใจ (Decision tree) หมายถึง ต้นไม้ที่ใช้ในการสนับสนุนการตัดสินใจ ซึ่งมีลักษณะเป็นโครงสร้างต้นไม้หัวกลับที่มีรากอยู่ด้านบนและใบอยู่ด้านล่างสุด โดยที่ภายในต้นไม้จะประกอบไปด้วยโหนด (Node) ซึ่งแต่ละโหนดนั้นจะแสดงถึงการตัดสินใจบนข้อมูลของคุณสมบัติต่างๆ กิ่งของต้นไม้แสดงถึงค่าหรือผลลัพธ์ที่ได้จากการทดสอบ และใบซึ่งเป็นสิ่งที่อยู่ด้านล่างสุดของต้นไม้ตัดสินใจจะแสดงถึงกลุ่มของข้อมูล (Class) หรือผลลัพธ์โหนดที่อยู่บนสุดจะเรียกว่าโหนดราก (Root Node)

1.8.3 Data Visualization เป็นการนำข้อมูลเชิงลึกที่ได้มาจากช่องทางต่าง ๆ มาวิเคราะห์และแสดงผลในรูปแบบของแผนภูมิกราฟรูปแบบที่หลากหลาย วิธีโอทีที่แสดงให้เห็นถึงการเปลี่ยนแปลงของข้อมูลเชิงปริมาณ ซึ่งจะทำให้ผู้อ่านเข้าใจได้ง่ายขึ้น ดูน่าสนใจมากกว่าการอ่านข้อมูลแบบตาราง

1.8.4 Web Browser หมายถึงโปรแกรมที่ใช้ในการเข้าถึงข้อมูลและติดต่อสื่อสารในรูปแบบ Webpage (เว็บไซต์) โดยโปรแกรมจะแปลงภาษาคอมพิวเตอร์ HTML ให้เป็นภาษาที่คนทั่วไปสามารถอ่านและเข้าใจได้บนหน้าเว็บ

1.8.5 CRISP-DM ย่อมาจาก Cross-industry standard process for data mining ซึ่งหมายถึง กระบวนการมาตรฐานที่ใช้สำหรับการทำเหมืองข้อมูล เพื่อทำการวิเคราะห์และนำไปใช้ประโยชน์ทางธุรกิจ มีอยู่ 6 ขั้นตอน

1.9 บทสรุป

จากบทนำที่ได้กล่าวมาในข้างต้นทั้งหมดนั้น ผู้วิเคราะห์ข้อมูลได้เล็งเห็นความสำคัญของการนำเทคโนโลยีมาประยุกต์ใช้ในการวิเคราะห์ข้อมูลนักศึกษาเบื้องต้น เพื่อเป็นตัวช่วยในการสนับสนุนการตัดสินใจลงทะเบียนเรียนและบอกถึงรายละเอียดของแต่ละรายวิชา เพื่อช่วยในการแก้ไขปัญหาจำนวนนักศึกษาที่ติดค้างหรือถอนรายวิชาด้วยกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM จากเทคนิคการทำเหมืองข้อมูลแบบการจัดกลุ่ม (Clustering) และแผนภูมิต้นไม้ (Decision tree) และเผยแพร่ข้อมูลสารสนเทศบน Web Browser ที่เป็นที่ยอมรับในยุคอินเทอร์เน็ตคือการเผยแพร่ทางสื่อออนไลน์ โดยใช้ภาษา HTML ในการพัฒนาเว็บไซต์สำหรับการเปิดเผยข้อมูลที่ได้มาจากหลักสูตรระบบสารสนเทศทางธุรกิจเนื่องจากเทคโนโลยีปัจจุบันเข้ามาช่วยในการจัดการข้อมูลหรือดำเนินงานต่าง ๆ ให้มีความสะดวกสบายเป็นอย่างมาก อีกทั้งผู้ใช้งานยังเข้าถึงข้อมูลได้ง่าย และมีประสิทธิภาพ